PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :
C07K 7/04, A23J 1/14
A61K 35/78, C12N 15/00, 9/12
C07H 15/12

A1

(11) International Publication Number: WO 87/ 02987

(43) International Publication Date: 21 May 1987 (21.05.87)

(21) International Application Number: PCT/US86/02444

(22) International Filing Date: 13 November 1986 (13.11.86)

(31) Priority Application Number: 798,163

(32) Priority Date: 13 November 1985 (13.11.85)

(33) Priority Country: US

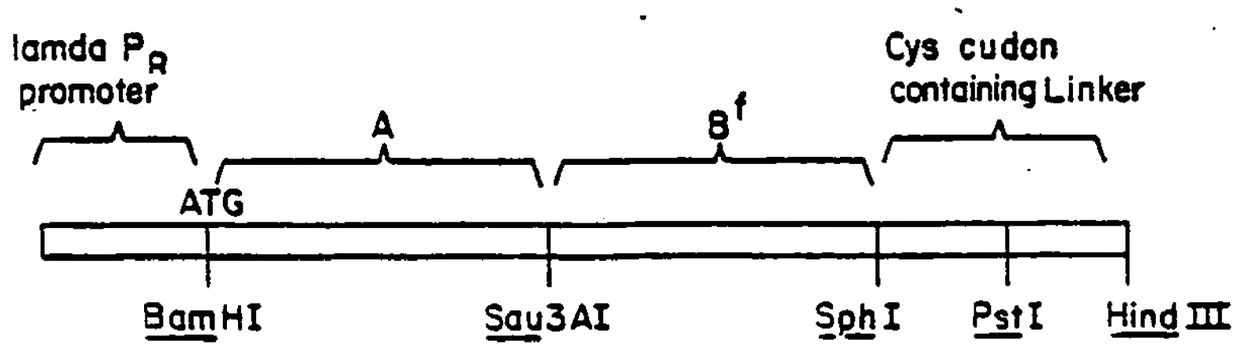
(71)(72) Applicant and Inventor: MURPHY, John, R. [US/US]; 32 Fifer Lane, Lexington, MA 02173 (US).

(74) Agent: CLARK, Paul, T.; Fish & Richardson, One Financial Center, Boston, MA 02111 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), NO, SE (European patent).

Published
With international search report.

(54) Title: CYS CODON-MODIFIED DNA



pABC1508

(57) Abstract

A DNA sequence encoding a fragment of a toxin molecule which is large enough to exhibit cytotoxic activity and small enough to fail to exhibit generalized eucaryotic cell binding, the DNA sequence including a non-naturally occurring cysteine codon.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**ML**	Mali
AU	Australia	**GA**	Gabon	**MR**	Mauritania
BB	Barbados	**GB**	United Kingdom	**MW**	Malawi
BE	Belgium	**HU**	Hungary	**NL**	Netherlands
BG	Bulgaria	**IT**	Italy	**NO**	Norway
BJ	Benin	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		

CYS CODON-MODIFIED DNA

Background of the Invention

This invention was made in part with government funding, and the government has certain rights in the invention.

This invention relates to the use of recombinant DNA techniques to make analogs of toxin molecules, and to the use of such molecules to treat medical disorders.

The literature contains a number of examples of hybrid molecules containing a specific binding ligand-portion and a toxin portion (e.g., ricin or diphtheria toxin); the ligand targets the toxin to an unwanted class of cells, sparing healthy cells, to which the ligand fails to bind.

For example, Bacha et al. U.S. Pat. No. 4,468,382 (hereby incorporated by reference) describes hybrid molecules made by derivatizing a neuropeptide hormone (e.g., thyrotropin releasing hormone) and an enzymically active fragment of diphtheria toxin using sulfur-containing groups and then reacting the derivatized molecules to join them via a disulfide bond. One disadvantage of this approach is that the site of derivatization on both molecules cannot be precisely controlled, so that the final product is heterogeneous, containing some molecules in which derivatization and coupling has impaired the toxicity or binding capacity of the hybrid molecule.

An approach which deals with this problem of heterogenicity is described in Murphy PCT International Publication No. WO/83/03971 (hereby incorporated by reference). The Murphy application describes hybrid proteins encoded by genes encoding both the toxin and

the specific binding portion of the hybrid protein. This approach, of course, can be used only for DNA-encoded peptide ligands.

Summary of the Invention

The present invention provides toxin molecules which can be linked to any specific-binding ligand, whether or not it is a peptide, at a position which is predeterminedly the same for every toxin molecule.

The invention generally features, in one aspect, a DNA sequence encoding a fragment of a toxin molecule which is large enough to exhibit cytotoxic activity and small enough to fail to exhibit generalized eucaryotic cell binding; the DNA sequence includes a non-naturally occuring cysteine codon, preferably located such that the fragment encoded by the DNA sequence, when linked to a cell-specific ligand via the cysteine residue encoded by the cysteine codon, exhibits cytotoxic enzymic activity.

In preferred embodiments, the toxin is diphtheria toxin, ricin, or abrin; the cysteine codon is introduced at the C-terminal-encoding end of the toxin-encoding DNA sequence or within 100 base pairs thereof; the ligand is a peptide hormone, a proteinaceous growth factor (preferably Interleukin I, Interleukin II, Interleukin III, or B-cell growth factor), an antibody, or a steroid hormone (e.g., estradiol).

In another aspect, the invention features a specific binding peptide ligand, and DNA sequences coding therefor, which can bind to any reactive sulfur group-containing toxin molecule in a predetermined and consistent manner. The DNA sequence of this aspect of the invention encodes a fragment of a ligand (preferably one of those listed above) which is large enough to

exhibit specific cell binding; the gene includes a non-naturally occuring cysteine codon, preferably located such that the fragment encoded by the DNA sequence, when linked to a toxin via the cysteine residue encoded by the cysteine codon, exhibits specific cell binding.

The invention also features the hybrid molecules made using the Cys-modified toxins and ligands of the invention, as well as the methods for making such hybrid molecules.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

Description of the Preferred Embodiments

The drawings will first briefly be described.

Drawings

Fig. 1 is a partial restriction map of a DNA fragment of the invention, in plasmid pABC1508.

Fig. 2 is a diagrammatic representation of the diphtheria toxin molecule.

Fig. 3 is a restriction map showing the location and orientatiqn of the diphtheria *tox* gene on the 3.9 *Bam*HI-I restriction fragment of corynephage beta^{tox}.

Figs. 4-5 are diagrammatic representations of the steps involved in the construction of pABC1508.

Fig. 6 is a diagrammatic representation of a plasmid, pMSH53 containing α-MSH-encoding DNA.

Fig. 7 is the nucleotide sequence of the *tox*^{228} allele and flanking regions, with amino acid residues shown above nucleotides; the *tox*^{228} allele is the same as the wild-type *tox* allele except for several mutations, notably the presence on the *tox*^{228} allele

of an NruI site (Fig. 7 was adapted from Fig. 1 of Kaczorek et al. (1983) Science 221, 855).

Tox Gene

The tox gene, and the diphtheria toxin molecule it encodes, will now briefly be described.

Figs. 2 and 3, illustrate, respectively, the diphtheria toxin molecule and the diphtheria tox gene, located on the 3.9 kb BamHI restriction fragment of corynephage beta^{tox}. Fig. 7 gives the sequence of the tox 228 allele.

Referring to Fig. 2, the diphtheria toxin molecule consists of several functional "domains" which can be characterized, starting at the amino terminal end of the molecule, as a hydrophobic signal sequence; enzymically active Fragment A, the fourteen amino acid exposed protease sensitive disulfide loop (DSL) l_1, containing a cleavage domain; Fragment B, which includes the lipid associating regions, e.g., a hydrophilic amphipathic domain and a hydrophobic domain; DSL l_2; and carboxy terminal end a. DSL l_1 contains three arginine residues; the Sau3A1 site between Fragment A and Fragment B (see Fig. 3) is at a position on the diphtheria toxin gene corresponding to the arginine residue farthest downstream of the three.

The process by which diphtheria toxin intoxicates sensitive eukaryotic cells involves at least the following steps: (i) diphtheria toxin binds to specific receptors on the surface of a sensitive cell; (ii) while bound to its receptor, the toxin molecule is internalized in an endocytic vesicle; (iii) either prior to internalization, or within the endocytic vesicle, the toxin molecule may be cleaved (or processed) at a site in the region of 47,000 daltons from the N-terminal end; (iv) as the pH of the endocytic vesicle decreases to

below 5.5, the processed form of toxin, while still bound to its receptor, spontaneously inserts into the endosomal membrane; (v) once embedded in the membrane, the lipid associating regions form a pore; (vi) a proteolytic cleavage in l_1, between Fragment A and B, occurs; (vii) thereafter, Fragment A, or a polypeptide containing Fragment A, is released into the cytosol; (viii) the catalytic activity of Fragment A, i.e., the nicotinamide adenine dinucleotide-dependent adenosine diphosphate ribosylation of Elongation Factor 2, causes the death of the intoxicated cell. It is apparent that a single molecule of Fragment A introduced into the cytosol is sufficient to kill a cell.

Modified Tox Gene

Referring to Fig. 1, there is shown the region of plasmid pABC1508 which encodes a peptide of the invention.

The DNA region shown in Fig. 1 includes the lambda P_R promoter (substituted for the promoter naturally associated with the tox gene); an ATG initiation site; a DNA sequence encoding enzymically active Fragment A of diphtheria toxin; a portion of the DNA region encoding Fragment B of diphtheria toxin; and a linker containing a Cys codon.

Referring to Figs. 1-3, the portion of the diphtheria tox gene used to make a DNA sequence of the invention includes the region encoding enzymically active Fragment A (and preferably the hydrophobic leader sequence preceding Fragment A), and a portion of the Fragment B-encoding region at least as long as that ending at the MspI site. As shown in Fig. 3, the Fragment A-encoding region (including the leader sequence) begins just downstream from a convenient Sau3AI site. The MspI site is the approximate location

of the end of the region of the tox gene which encodes cross reacting material 45 (CRM 45), described in Bacha et al., id. This portion of the diphtheria toxin molecule contains the lipid associating regions of Fragment B, but does not contain l_2, and is represented in Fig. 2 as the portion of Fragment B between y and z. The Fragment B-encoding region employed can end anywhere beyond MspI, up to the SphI site. If the SphI site is used, l_2 is included, and the portion of Fragment B is that between y and x in Fig. 3. As previously mentioned, any region ending between MspI and SphI can be used; one example is the region ending at the position of NruI, which, like the region ending at MspI, encodes a Fragment which does not contain l_2. Any region shorter than one ending at MspI should not be used because such a fragment will not include enough of the transverse lipid associating region and thus not bring about pore formation, which is necessary for toxicity. Portions of Fragment B encoded by regions ending downstream of SphI should not be used to avoid including the diphtheria toxin receptor binding domain. (NruI is not found on the wild-type tox allele, but only on the mutant tox^{228} allele, described in Kaczorek et al. (1983) Science 221, 855.)

In the illustrated DNA construct (Fig. 1) the Cys codon is located at the C-terminal end of the tox-encoding DNA sequence. This location ensures that the linker containing the Cys codon will not interfere with the enzymic activity of Fragment A. Other locations in the molecule which are downstream from the Fragment A encoding region can also be used, i.e., the Cys codon-containing linker can be inserted anywhere in the Fragment B-encoding region.

Other toxins which are DNA- encoded amino acid chains can be used, in addition to diphtheria toxin; examples are ricin and the plant toxin abrin.

Ligands

The specific-binding ligands used in the invention can consist of an entire ligand, or a portion of a ligand which includes the entire binding domain of the ligand, or an effective portion of the binding domain. It is most desirable to include all or most of the binding domain of the ligand molecule. In the case of alpha-MSH, a small peptide of thirteen amino acids, or beta-MSH, which contains seventeen amino acids, the portion of the molecule consisting of nine amino acids at the carboxy terminal end of the molecule, which contains the receptor-specific binding domain, can be used, or, more preferably, the entire molecule can be used. It is most preferred that at least a portion of the ligand not involved in cell binding be included, so that derivatization can be carried out in this nonbinding portion, minimizing the chance that derivatization will interfere with binding. For example, derivatization of alpha-MSH is preferably carried out at or near the N-terminal end of the molecule, because the C-terminal end contains the specific binding domain.

The regions within cell-specific ligands in which the binding domain is located are now known for a number of such ligands. Furthermore, recent advances in solid phase polypeptide synthesis can enable those skilled in this technology to determine the binding domain of practically any peptide ligand, by synthesizing various fragments of the ligand, and testing them for the ability to bind the class of cells to be killed.

The specific class of cells which are bound and killed by the hybrids of the invention is determined by the specific ligand which imparts the binding domain of the hybrid molecule. Any cell-specific ligand can be used which has a binding domain which is specific for a particular class of cells which are to be killed. Polypeptide hormones are useful such ligands. Hybrid proteins made using alpha- or beta-MSH, for example, can selectively bind to melanocytes, rendering the hybrids useful in the treatment of primary melanoma and metastic melanoma loci. Other specific-binding ligands which can be used include the proteinaceous growth factors interleukin I, interleukin II, interleukin III, and B-cell growth factor. Interleukin II is of particular importance because of its role in allergic reactions and autoimmune diseases such as Systemic Lupus Erythmatosis (SLE), involving activated T cells. Hybrids made using B-cell growth factor can be used as immunosuppressant reagents which kill proliferating B-cells, which bear B-cell growth factor receptors, and which are involved in hypersensitivity reactions and organ rejection.

The other major class of specific binding proteins are antibodies. The antibodies most useful are those against tumors; such antibodies (generally monoclonal) are already well-known targeting agents used in conjunction with covalently bound cytotoxins. In the present invention, the anti-tumor antibodies (preferably not the whole antibody, but just the Fab portion) are those which recognize a surface determinant on the tumor cells and are internalized in those cells via receptor-mediated endocytosis; antibodies which are capped and shed will not be as effective.

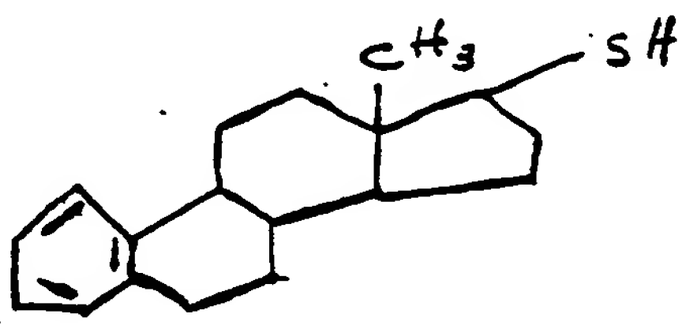
Other useful polypeptide ligands having cell-specific binding domains are somotostatin, follicle

stimulating hormone (specific for ovarian cells); luteinizing hormone (specific for ovarian cells); thyroid stimulating hormone (specific for thyroid cells); vasopressin (specific for uterine cells, as well as bladder and intestinal cells); prolactin (specific for breast cells); and growth hormone (specific for certain bones cells).

Peptide hormones must be derivatized with a sulfhydryl group reactive with the Cys of the toxin molecule. This can be carried out by inserting a Cys codon-containing linker into an appropriate location in a DNA sequence encoding the hormone, in a manner analogous to that described below for the tox gene. Alternatively, a sulfhydryl group, either by itself or as part of a Cys residue, can be introduced using solid phase peptide synthesis techniques. For example, the introduction of sulfhydryl groups into peptides is described in Hiskey (1981) Peptides 3, 137. Derivatization can also be carried out according to the method described for the derivatization of the peptide hormone thyrotropin releasing hormone in Bacha et al. U.S. Pat. No. 4,468,382, id. Similarly, proteins can be derivatized at the DNA or protein chemistry level. The introduction of sulfhydryl groups into proteins is described in Maasen et al. (1983) Eur. J. Biochem. 134, 2, 32.

The major class of non-peptide specific binding ligands useful in the invention are the steroid hormones. One example is estrogen and estrogen derivatives such as estradiol; these are currently used in the treatment of prostate carcinoma and post-menopausal mammary carcinoma. Hybrids containing these hormones, or analogs thereof, can be used in the same or smaller dosages, to treat the same diseases.

The derivatization of steroid hormones can be carried out using standard techniques, such as those described in Ikagawa et al. (1981) J. Org. Chem. 46, 18, 3747. For example, the steroid hormone 17-beta-estradiol can be derivatized, substituting an SH group for a hydroxyl group, to yield



Gene Construction

Generally, plasmids are manipulated according to standard techniques. Plasmid DNA is digested with restriction endonucleases as recommended by the manufacturer (e.g., New England Biolabs, Beverly, Mass.). Restriction fragments are electrophoresed in 1% horizontal agarose gels for 30-60 minutes at 80-100 V in TBE (89 mM boric acid, 89 mM Trizma base [Sigma Chemical Co., St. Louis, Mo.], 2.5 mM EDTA, pH 7.0) in the presence of 200 ng/ml ethidium bromide. Small DNA fragments are electrophoresed in 8% vertical polyacrylamide gels at 100 V for 2-5 hours, and stained with ethidium bromide. Gels are photographed on an ultraviolet transilluminator on Polaroid type 667 film using a red filter.

Plasmid pABC508 was constructed by fusing two pieces of DNA, one encoding Fragment A, and the other encoding part of Fragment B, to which a Cys codon-containing linker had been attached.

Referring to Fig. 4, this fusion was constructed from two plasmids, pDT201, which contains the fragment A-encoding region, and pDT301, which

contains most of the fragment B-encoding region of the diphtheria toxin gene. The construction of each of these pieces of DNA is described below.

Plasmid pDT301 was constructed by cutting out of the tox allele a Sau3AI-1 sequence encoding all but the C-terminal 17 amino acids of Fragment B. This sequence, which carries the restriction endonuclease sites ClaI, MspI, and SphI, was inserted into the BamHI site of plasmid pUC8 (described in Viera et al. (1982) Gene 19, 259) to yield pDT301. Plasmid pDT201 contains the Fragment A-encoding Sau3AI-2 sequence (Fig. 3) (see Leong et al. (1983) Science 220, 515). (pDT301 and pDT201, in E. coli, have been deposited in the American Type Culture Collection, Rockville, MD and given ATCC Accession Nos., respectively, 39360 and 39359. Applicant's licensee, Seragen, Inc., acknowledges its responsibility to replace these cultures should they die before the end of the term of a patent issued hereon, and its responsibility to notify the ATCC of the issuance of such a patent, at which time the deposits will be made available to the public for a period of at least 30 years after the date of deposit. Until that time the deposits will be made available to the Commissioner of Patents under the terms of 37 CFR §1.14 and 35 USC §112.)

Still referring to Fig. 4, plasmid pDT301 was modified by the addition of a Cys codon-containing linker as follows. A synthetic linker was constructed on a controlled pore glass solid phase support in a 380A DNA Synthesizer (Applied Biosystems, Inc., Foster City, CA) by hybridization of 21-mer and 29-mer oligonucleotides through a 21 bp homologous core, leaving a 4bp 1/2 SphI and 1/2 HindIII single-stranded sequence on each end. This linker has the sequence

AlaAlaAlaCysStp
5'-CGGCTGCAGCATGTTAGTAGA-3'
3'-GTACGCCGACGTCGTACAATGATCTTCGA-5'
1/2 SphI PstI C 1/2 HindIII

The linker encodes three alanine residues, and contains a Cys codon (TGT) and a Stop codon (TAG). (This design not only allows for the expression of a Cys-containing peptide according to the present invention, but also could allow for the insertion at the PstI site of a gene encoding a specific binding ligand.)

As shown in Fig. 4, pDT301 was digested with SphI and HindIII to remove the DNA region designed "E" in Fig. 4, and the Cys codon-containing linker was then ligated into the plasmid at the SphI, HindIII sites to give plasmid pBC508. pBC508 was then cut with HindIII and Sau3AI to give Fragment 1.

Still referring to Fig. 4, plasmid pDT201 was digested with HindIII and the single-stranded ends filled in with DNA polymerase I (Klenow fragment). The resulting blunt ends were
ligated to the double-stranded EcoRI linkers CCTTAAGG(New
GGAATTCC

England Biolabs, Beverly, MA) to give pDT201', which was then cut with EcoRI and Sau3A to give Fragment 2.

Fragments 1 and 2 were mixed in equimolar concentrations ligated together, according to standard procedures, and the mixture was then digested with EcoRI and HindIII. The digested mixture was then ligated into the EcoRI and HindIII digested pEMBL8 (Dente et al. (1983) Nucleic Acid Res. 11, 1645), which contains unique EcoRI and HindIII sites, to give pABC508. Plasmid pABC508 can be transformed into a suitable host, e.g., E. coli, as described below, to produce Cys-modified toxin molecules.

Alternatively, the naturally occurring tox promoter can be replaced with a different promoter, as follows.

The lambda P_R promoter is contained in the expression vector pEMBL8ex3 (Dente et al., id). Referring to Fig. 5, the DNA sequence around the initiation site of the tox gene is shown, as are the corresponding amino acids. pABC508 was cut with EcoRI and then treated with Bal31 for a period of 10-15 minutes at 37°C with one unit of enzyme per microgram of DNA. The resulting mixture of DNA fragments was ligated to

the BamHI linkers CCTAGGCC, transformed into E. coli HB101
GGATCCGG

(Bethesda Laboratories, Gaithersburg, MD), and the DNA sequence of the region encoding the 5' end of tox, and the sequence of 30 of the resulting clones determined. One clone, containing the DNA sequence shown in Fig. 5, was purified and the BamHI-HindIII fragment isolated and inserted into pEMBL8ex3 which had been cut with BamHI and HindIII. The resulting plasmid, pABC1508, contains the lambda P_R promoter and an ATG translational start codon. An extra asparagine and proline residue are inserted during this process. In Fig. 5, Cro represents the Cro gene of lambda, and SD represents the Shine-Dalgarno sequence. The lambda P_R promoter can be regulated by the lambda cI gene. In this example the mutant cI_{857} temperature-sensitive repressor gene is used such that the P_R promoter is inactive at 30°C and active at 37°C.

pABC1508 was transformed, using conventional techniques (e.g., as described in Maniatis et al. (1984) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, N.Y.), into E. coli HB101 (others, e.g., E. coli

JM101 or 5Y327, can also be used) and the expression of the diphtheria tox gene products analyzed. The introduction of the positively charged asparagine residue in the tox signal sequence does not affect the export of the tox polypeptides into the periplasmic compartment of the recombinant host.

E. coli cells transformed with vectors containing Cys-modified toxin-excoding DNA are grown under standard culture conditions, e.g., in Luria Broth containing, per liter, 10 g tryptone, 10 g NaCl, and 5 g yeast extract, and supplemented with 100 g/ml ampicillin. The diphtheria toxin-related molecules, which are exported to the periplasmic space, are purified from periplasmic extracts. Periplasmic extracts are prepared from cells grown in 9.5 liter volumes at 37°C to an A_{590} of approximately 1.0. If the natural tox promoter has been replaced with temperature sensitive cI857 regulatory sequences under the control of the temperature-sensitive cI_{857} gene, as described herein, cells are grown at 30°C, and expression is induced by increasing the incubation temperature to 42°C for 15 min. The culture is then grown at 40°C for an additional hour. In either instance, the culture is concentrated to approximately 1 liter by filtration through 0.45 μ membranes (Pellicon system, Millipore Corp., Bedford, Mass.) and chilled to -4°C. Bacteria are harvested by centrifugation, resuspended in ice cold 20% sucrose, 30mM Tris-HCl, 1 mM EDTA, pH 7.5, and then digested with lysozyme (750 g/ml final concentration) for 30 minutes. Spheroplasts are removed by centrifugation, 2 mg p-amidinophenylmethylsulfonylfluoride (p-APMSF,

Calbiochem, San Diego, Calif.) is added, and the periplasmic extract is sterilized by filtration through 0.2 μ membranes.

The Cys-modified toxin-related molecules are then purified by chromatography on Phenyl-Sepharose (Pharmacio Fine Chemicals, Piscataway, N.J.) and DEAE-cellulose essentially as described by Rappuoli et al. (1985) Biotechnology, p. 165. Periplasmic extracts are dialysed against 10mM sodium phosphate (pH 7.2) buffer, and ammonium sulfate added to 13% (w/v). The crude extracts are then applied to a Phenyl-Sepharose column equilibrated with 10 mM phosphate buffer containing 13% ammonium sulfate. The modified toxin is eluted and dialysed against 10 mM phosphate buffer, and then applied to DEAE-cellulose column. After washing with phosphate buffer, the DEAE-cellulose column is developed with a linear NaCl gradient in phosphate buffer.

The modified toxin is then applied to an anti-diphtheria toxin immunoaffinity column, containing antibody made as described in Zucker et al. (1984) Molecular Immunol. 21, 785. Following extensive wahsing, the modified toxin is eluted with 4 M guanidine hydrochloride, and immediately dialysed against phosphate buffer. The purified modified toxin is then concentrated to approximately 100 g/ml by placing the dialysis bag in dry Sephadex G-200. All purification procedures are carried out at 4°C, and the modified toxin is stored in small aliquots at -76°C until used.

Specific Binding Ligand: Alpha-MSH

Referring to Fig. 6, there is shown plasmid pMSH53, which contains a DNA insert encoding alpha-MSH. pMSH53 was made by inserting into pUC8 an alpha-MSH-

encoding sequence having a 1/2 PstI site at each end:

```
   5' GCAAGTTATAGCATGGAACATTTTAGATGGGGAAAACCTGTATAGCTGCA 3'
3' ACGTCGTTCAATATCGTACCTTGTAAAATCTACCCCTTTTGGACATATCG 5'
1/2 PstI                                                1/2 PstI
```

Plasmid pMSH53 can be used to create, using conventional methods, an expression vector for the production of alpha-MSH in cultured bacterial cells e.g., E. coli. Further, prior to such expression, a Cys-containing linker can be fused to or near the N-terminal end of the alpha-MSH-encoding sequence, in a manner analogous to the method described above for the tox gene, to produce a modified alpha-MSH containing an N-terminal Cys capable of reacting with the added Cys of the diphtheria toxin fragment. Alternatively, the Cys-containing MSH molecule can be chemically linked to any toxin molecule on which a reactive sulfhydryl group is present or has been added post-translationally, i.e., at the protein chemistry, not the DNA, level.

Rather than making alpha-MSH using recombinant DNA techniques as described above, alpha-MSH can be purified from biological sources, or obtained commercially (e.g., from Sigma Chemical Co., St. Louis, MO).

Chemical Linkage

After an available Cys has been added to the ligand by genetic engineering techniques, or if the ligand contains an available reactive Cys or other sulfur-containing group, the toxin and ligand are coupled by reducing both compounds and mixing toxin and ligand, in a ratio of about 1:5 to 1:20, and the disulfide reaction is allowed to proceed at room

temperature to completion (generally, 20 to 30 minutes). The mixture is then dialyzed extensively against phosphate buffered saline to remove unreacted ligand molecules. The final purification step involves the separation, on the basis of size, of the desired toxin-hormone conjugates from toxin-toxin and hormone-hormone dimers; this is done by carrying out, in phosphate-buffered saline, Sephadex G100 chromotography.

Use

The hybrid toxin-ligand molecules are administered to a mammal, e.g., a human, suffering from a medical disorder, e.g., cancer, characterized by the presence of a class of unwanted cells to which the ligand can selectively bind. The amount of hybrid molecule administered will vary with the type of disease, extensiveness of the disease, and size and species of the mammal suffering from the disease. Generally, amounts will be in the range of those used for other cytotoxic agents used in the treatment of cancer, although in certain instances lower amounts will be needed because of the specificity of the molecules.

The hybrid molecules can be administered using any conventional method; e.g., via injection, or via a timed-release implant. The hybrid proteins can be combined with any non-toxic, pharmaceutically-acceptable carrier substance.

In the case of MSH hybrids, topical creams can be used to kill primary melanoma cells, and injections or implants can be used to kill metastatic cells.

Estradiol hybrids, exhibiting binding specificity for certain breast cancers characterized by cells bearing estradiol receptors, can be used to treat primary and metastatic cells.

Other embodiments are within the following claims.

Claims

1. A targeted toxin molecule comprising a toxic portion comprising a toxin molecule which is large enough to exhibit cytotoxic activity and small enough to fail to exhibit generalized eucaryotic cell binding, said toxic portion including a non-naturally occuring cysteine and being encoded by a DNA sequence, said toxic portion being chemically linked by said cysteine to a cell specific ligand comprising a peptide, a proteinaceous growth factor, or a steroid hormone.

2. The tartgeted toxin molecule of claim 1 wherein said cysteine is located such that said toxin molecule, when linked to said cell-specific ligand via said cysteine residue, exhibits cytotoxic enzymic activity.

3. The targeted toxin molecule of claim 1 wherein said toxin is diphtheria toxin, ricin, or abrin.

4. The targeted toxin molecule of claim 1 wherein said toxic portion and said cell specific ligand are linked by a disulfide linkage.

5. The targeted toxin molecule of claim 4 wherein said ligand includes a sulfhydryl group and said disulfide linkage is between said cysteine of said toxin molecule and the sulfhydryl group of said ligand.

6. The targeted toxin molecule of claim 1 wherein said peptide is a hormone.

7. The targeted toxin molecule of claim 1 wherein said proteinaceous growth factor is Interleukin I, Interleukin II, Interleukin III, or B-cell growth factor.

8. The targeted toxin molecule of claim 1 wherein said steroid hormone is estradiol.

9. A DNA sequence encoding a fragment of a ligand which is large enough to exhibit specific cell binding, said DNA sequence including a non-naturally occuring cysteine codon.

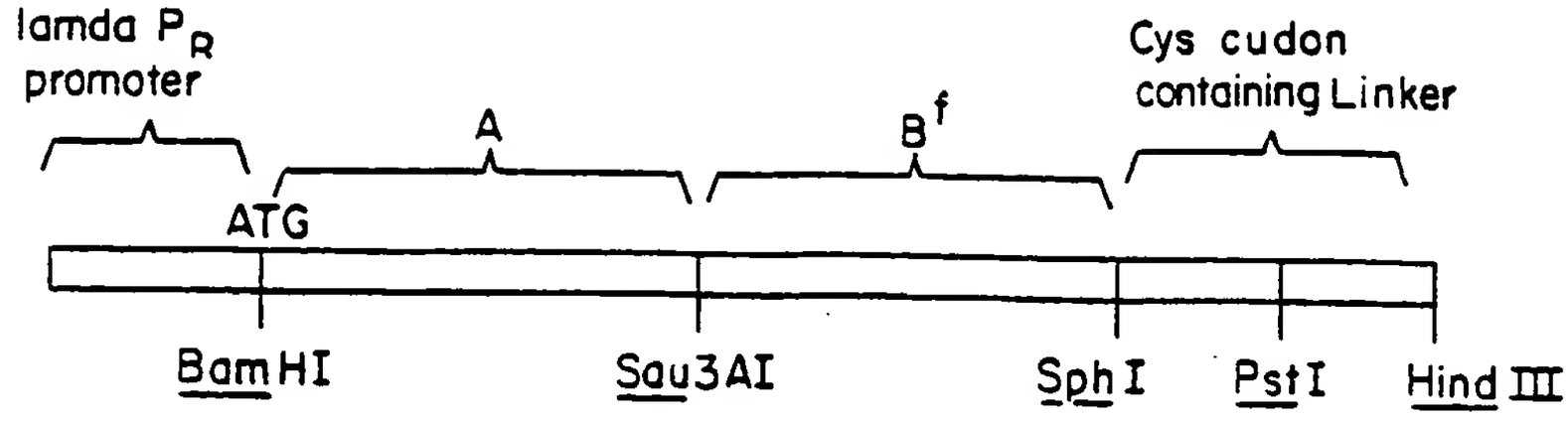
10. The DNA sequence of claim 9 wherein said cysteine codon is located such that said fragment encoded by said DNA sequence, when linked to a toxin molecule via the cysteine residue encoded by said cysteine codon, exhibits specific cell binding.

11. The DNA sequence of claim 9 wherein said ligand is a peptide, a proteinaceous growth factor, or an antibody.

12. The DNA sequence of claim 11 wherein said peptide is a hormone.

13. The DNA sequence of claim 11 wherein said proteinaceous growth factor is Interleukin I, Interleukin II, Interleukin III, or B-cell growth factor.

14. The cell specific ligand encoded by the DNA sequence of claim 9.



pABC1508

FIG 1

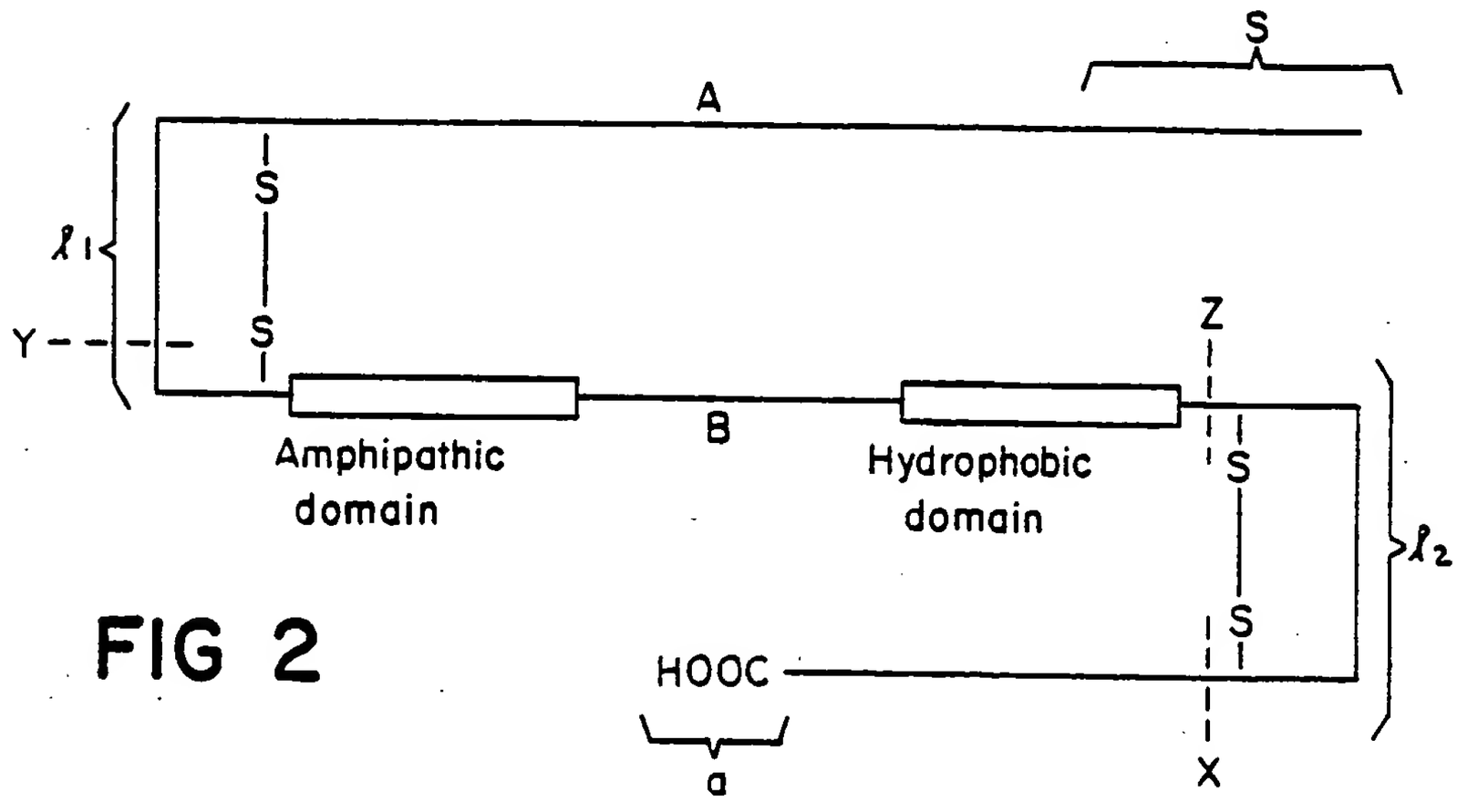


FIG 2

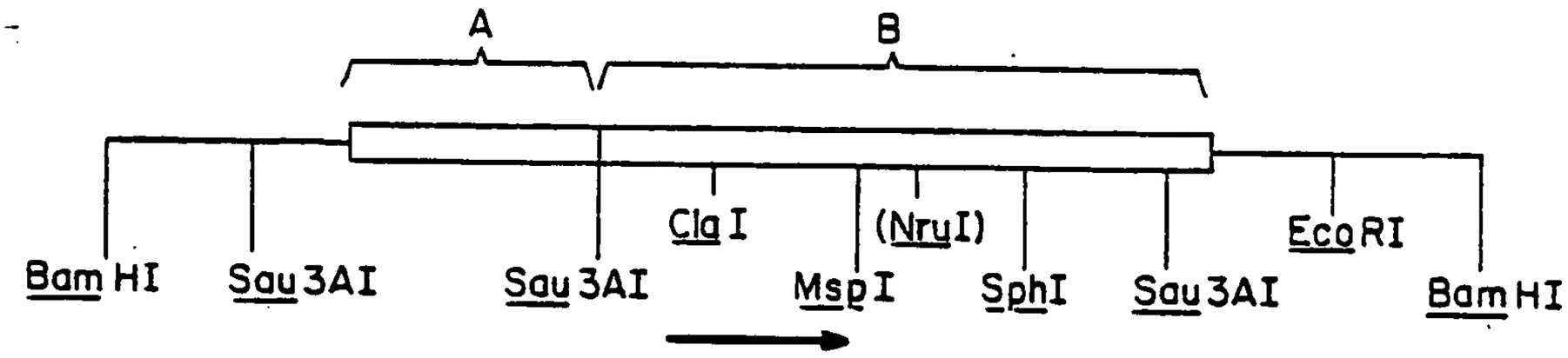


FIG 3

pABC 508
pEMBL8
DIGESTED
FRAGMENT 1
HindIII
PstI
SphI
Sau3AI
HindIII, Sau3AI DIGEST
pBC508
HindIII
PstI
SphI
Sau3AI
SphI, HindIII DIGEST
Linker
pDT301
HindIII
PstI
SphI
SauI
(NruI)
MspI
ClaI
SauI
EcoRI
(1) LIGATE
(2) EcoRI, HindIII DIGEST
(3) LIGATE INTO EcoRI and HindIII — DIGESTED pEMBL8
FRAGMENT 2
pDT201'
EcoRI, Sau3AI DIGEST
pDT201
Linker
EcoRI
HindIII DIGEST
Sau3AI
HindIII

FIG 4

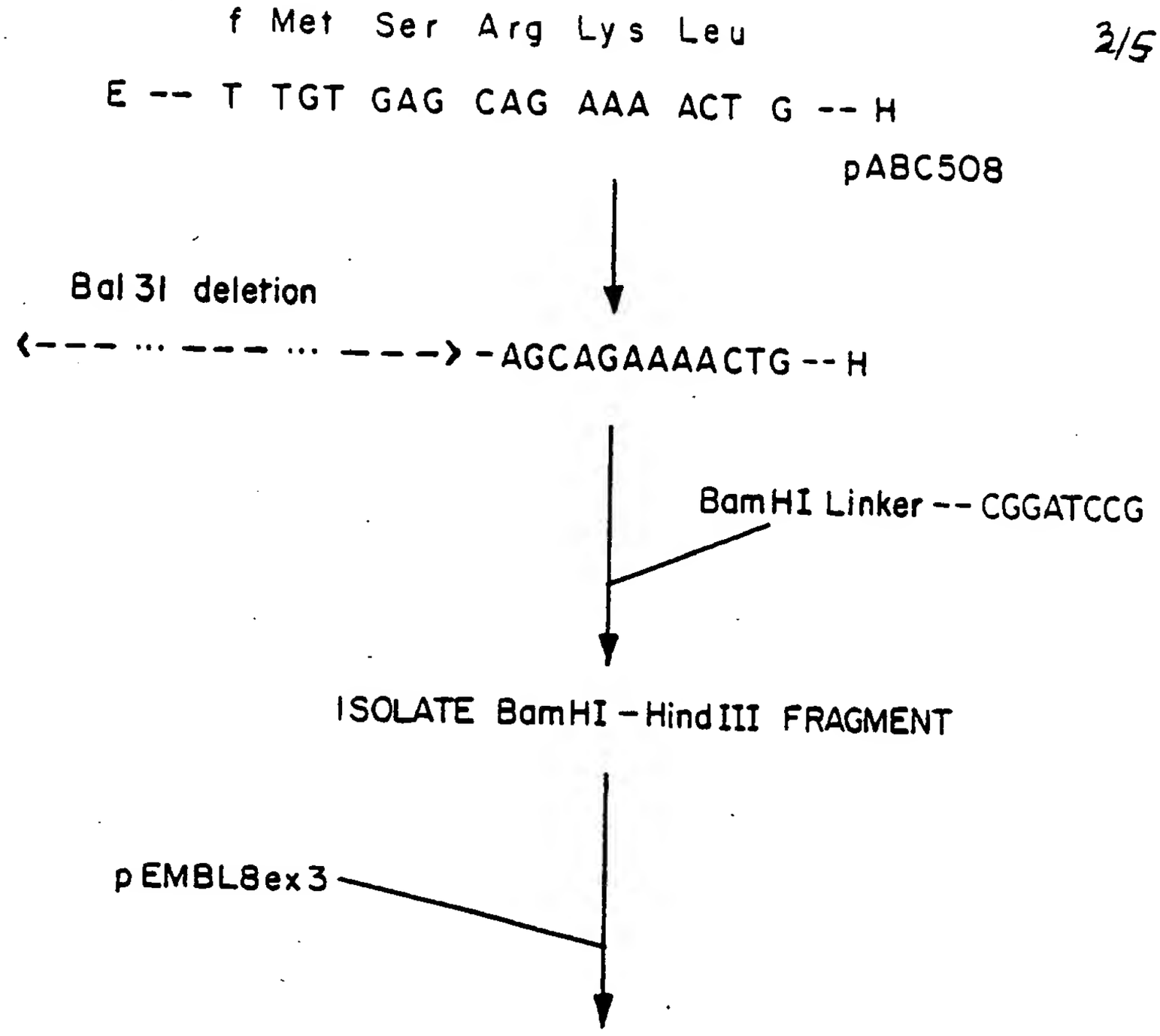


FIG 5

pABC1508

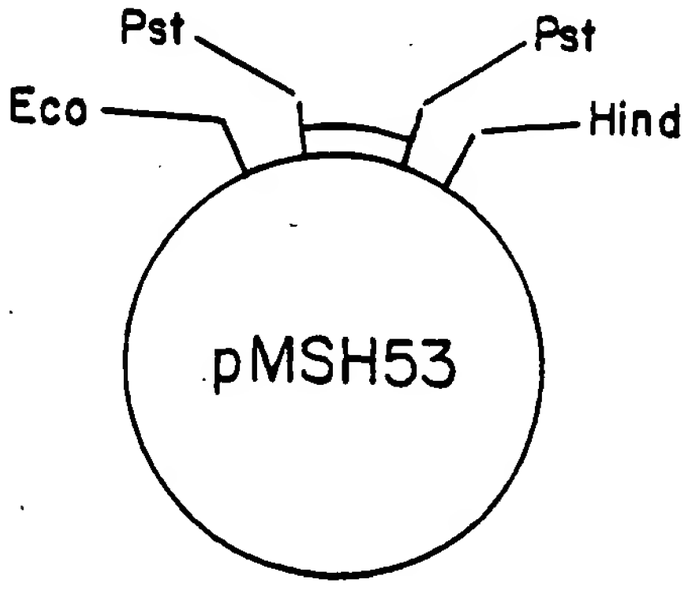


FIG 6

4/5

FIG. 7A

```
ATCTTTCCGG TGTGGTACAC CTGATCTGGT CCGGTTCATG TTGTGGTGGT CAACGCTGGG
-300                                                    -250

GTAACCGGCG TTGCGTATCC AGTGGCTACA CTCAGGTTGT AATGATTGGG ATGATGTACC TGATCTGAGA GCGATTAAAA ACTCATTGAG GAGTAGGTCC CGATTGGTTT TTGCTAGTGA
                                            -200                                                 -150

AGCTTAGCTA GCTTTCCCCA TGTAACCAAT CTATCAAAAA AGGGCATTGA TTTCAGAGCA CCCTTATAAT TAGGATAGCT TTACCTAATT ATTTTATGAG TCCTGGTAAC GGGATACCTT
                -100                                                     -50                                        SD        -1

S->                            10                                          20                   A->             30
fMET SER ARG LYS LEU PHE ALA SER ILE LEU ILE GLY ALA LEU LEU GLY ILE GLY ALA PRO PRO SER ALA HIS ALA|GLY ALA ASP ASP VAL VAL ASP SER
 GTG AGC AGA AAA CTG TTT GCG TCA ATC TTA ATA GGG GCG CTA CTG GGG ATA GGG GCC CCA CCT TCA GCC CAT GCA|GGC GCT GAT GAT GTT GTT GAT TCT

                        40                     50                                    60
SER LYS SER PHE VAL MET GLU ASN PHE SER SER TYR HIS GLY THR LYS PRO GLY TYR VAL ASP SER ILE GLN LYS GLY ILE GLN LYS PRO LYS SER GLY
TCT AAA TCT TTT GTG ATG GAA AAC TTT TCT TCG TAC CAC GGG ACT AAA CCT GGT TAT GTA GAT TCC ATT CAA AAA GGT ATA CAA AAG CCA AAA TCT CGT
100

            70                         80                                    90
THR GLN GLY ASN TYR ASP ASP ASP TRP LYS GLY PHE TYR SER THR ASP ASN LYS TYR ASP ALA ALA GLY TYR SER VAL ASP ASN GLU ASN PRO LEU SER
ACA CAA GGA AAT TAT GAC GAT GAT TGG AAA GCG TTT TAT AGT ACC GAC AAT AAA TAC GAC GCT GCG GGA TAC TCT GTA GAT AAC GAA AAC CCG CTC TCT
200

100             GLY                 110                             120                                    130
GLY LYS ALA GLY (ASP) VAL VAL LYS VAL THR TYR PRO GLY LEU THR LYS VAL LEU ALA LEU LYS VAL ASP ASN ALA GLU THR ILE LYS LYS GLU LEU GLY
GGA AAA GCT GGA (GAC) GTG GTC AAA GTG ACG TAT CCA GGA CTG ACG AAG GTC CTC GCA CTA AAA GTG GAT AAT GCC GAA ACT ATT AAG AAA GAG TTA CGT
300

                        140                           150                                    160
LEU SER LEU THR GLU PRO LEU MET GLU GLN VAL GLY THR GLU GLU PHE ILE LYS ARG PHE GLY ASP GLY ALA SER ARG VAL VAL LEU SER LEU PRO PHE
TTA AGT CTC ACT GAA CCC TTG ATG GAG CAA GTC GGA ACA GAA GAG TTT ATC AAA AGG TTC GGG GAT GGT GCT TCG CGT GTA GTG CTC AGC CTT CCC TTT
400

            170  ▽      ▼                   180                  GLU            190
ALA GLU GLY SER SER SER VAL GLU TYR ILE ASN ASN TRP GLU GLN ALA LYS ALA LEU SER VAL (LYS) LEU GLU ILE ASN PHE GLU THR ARG GLY LYS ARG
GCC GAG GGG AGT TCT AGC GTC GAA TAT ATT AAT AAC TGG GAA CAG GCG AAA GCG TTA AGC GTA (AAA) CTT GAG ATT AAT TTT GAA ACC CGT GGA AAA CGT
500

200                              210                              B|--220     SER          LYS                    230
GLY GLN ASP ALA MET TYR GLU TYR MET ALA GLN ALA CYS ALA GLY ASN ARG VAL ARG ARG|SER VAL GLY (GLY) SER LEU (SER) CYS ILE ASN LEU ASP TRP
GGC CAA GAT GCG ATG TAT GAG TAT ATG GCT CAA GCC TCT GCA GGA AAT CGT GTC AGG CGA|TCA GTA GGT (GGC) TCA TTG (TCA) TGC ATA AAT CTT GAT TGG
600

                        240                                250 Sau 3 AI                      260
ASP VAL ILE ARG ASP LYS THR LYS THR LYS ILE GLU SER LEU LYS GLU HIS GLY PRO ILE LYS ASN LYS MET SER GLU SER PRO ASN LYS THR VAL SER
GAT GTC ATA AGG GAT AAA ACT AAG ACA AAG ATA GAG TCT TTG AAA GAG CAT GGC CCT ATC AAA AAT AAA ATG AGC GAA AGT CCC AAT AAA ACA GTA TCT
700
```

FIG. 7B

```
                                270                                                 280                                                 290
GLU GLU LYS ALA LYS GLN TYR LEU GLU GLU PHE HIS GLN THR ALA LEU GLU HIS PRO GLU LEU SER GLU LEU LYS THR VAL THR GLY THR ASN PRO VAL
GAG GAA AAA GCT AAA CAA TAC CTA GAA GAA TTT CAT CAA ACG GCA TTA GAG CAT CCT GAA TTG TCA GAA CTT AAA ACC GTT ACT GGG ACC AAT CCT GTA
          800
        300                                                 310                                                 320                                                 330
PHE ALA GLY ALA ASN TYR ALA ALA TRP ALA VAL ASN VAL ALA GLN VAL ILE ASP SER GLU THR ALA ASP ASN LEU GLU LYS THR THR ALA ALA LEU SER
TTC GCT GGG GCT AAC TAT GCG GCG TGG GCA GTA AAC GTT GCG CAA GTT ATC GAT AGC GAA ACA GCT GAT AAT TTG GAA AAG ACA ACT GCT GCT CTT TCG
          900
                                340                                                 350                                                 360
ILE LEU PRO GLY ILE GLY SER VAL MET GLY ILE ALA ASP GLY ALA VAL HIS HIS ASN THR GLU GLU ILE VAL ALA GLN SER ILE ALA LEU SER SER LEU
ATA CTT CCT GGT ATC GGT AGC GTA ATG GGC ATT GCA GAC GGT GCC GTT CAC CAC AAT ACA GAA GAG ATA GTG GCA CAA TCA ATA GCT TTA TCG TCT TTA
                1000
                370                                                 380                                                 390
MET VAL ALA GLN ALA ILE PRO LEU VAL GLY GLU LEU VAL ASP ILE GLY PHE ALA ALA TYR ASN PHE VAL GLU SER ILE ILE ASN LEU PHE GLN VAL VAL
ATG GTT GCT CAA GCT ATT CCA TTG GTA GGA GAG CTA GTT GAT ATT GGT TTC GCT GCA TAT AAT TTT GTA GAG AGT ATT ATC AAT TTG TTT CAA GTA GTT
              1100
        400                                                 410   CB44 >    ASN                                             420
HIS ASN SER TYR ASN ARG SER ALA TYR SER PRO GLY HIS LYS THR GLN PRO PHE LEU HIS ASP GLY TYR ALS VAL SER TRP ASN THR VAL GLU ASP SER
CAT AAT TCG TAT AAT CGT TCC GCC TAT TCT CCG GGC CAT AAA ACG CAA CCA TTT CTT CAT GAC GGG TAT GCT GTC AGT TGG AAC ACT GTT GAA GAT TCG
              1200
430                                 440                                                 450                 NruI    GLY             460
ILE ILE ARG THR GLY PHE GLN GLY GLU SER GLY HIS ASP ILE LYS ILE THR ALA GLU ASN THR PRO LEU PRO ILE ALA SER VAL LEU LEU PRO THR ILE
ATA ATC CGA ACT GGT TTT CAA GGG GAG ACT GGG CAC GAC ATT AAA ATT ACT GCT GAA AAT ACC CCG CTT CCA ATC GCG AGT GTC CTA CTA CCG ACT ATT
              1300
                    470                                                 480                                                 490
PRO GLY LYS LEU ASP VAL ASN LYS SER LYS THR HIS ILE SER VAL ASN GLY ARG LYS ILE ARG MET ARG CYS ARG ALA ILE ASP GLY ASP VAL THR PHE
CCT GGA AAG CTG GAC GTT AAT AAG TCC AAG ACT CAT ATT TCC GTA AAT GGT CGA AAA ATA AGG ATG CGT TGC AGA GCT ATA GAC GGT GAT GTA ACT TTT
              1400
        500                                                 510                                                 520
CYS ARG PRO LYS SER PRO VAL TYR VAL GLY ASN GLY VAL HIS ALA ASN LEU HIS VAL ALA PHE HIS ARG SER SER SER GLU LYS ILE HIS SER ASN GLU
TGT CGC CCT AAA TCT CCT GTT TAT GTT GGT AAT GGT GTG CAT GCG AAT CTT CAC GTG GCA TTT CAC AGA AGC AGC TCG GAG AAA ATT CAT TCT AAC CAA
              1500
    530     ASN                                             540                                                 550     LYS             560
ILE SER SER ASP SER ILE GLY VAL LEU GLY TYR GLN LYS THR VAL ASP HIS THR LYS VAL ASN SER LYS LEU SER LEU PHE PHE GLU ILE LYS SER ...
ATT TCG TCG GAT TCC ATA GGC GTT CTT GGG TAC CAG AAA ACA GTA GAT CAC ACC AAG GTT AAT TCT AAG CTA TCG CTA TTT TTT GAA ATC AAA AGC TCA
              1600

AAGGTAG TGGGGTCGTG TGCCGGTAAG CGGAACGGTT CCGGAATGGC GCTATAGTAT GCACAGGTAG AGCAGAATTC GAATCTGACT ACGGATCAGA AGGTTGGGGG TTCGAATCCC TCC
                    1700                                                                                                                1800

GGGCGCA CAAGTGAAAC CCCAGCTCAT AGCATGTTTG AGCTGGGGTT TCTCATGGCG TGTGGGTTGT CTGACTGTTG GCTGTTGTTG CGGTGGTTGG TGCTCGTACC GAACCGAACG
                                                                                                                        1900
```

INTERNATIONAL SEARCH REPORT

International Application No PCT/US86/02444

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) 3

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(4): C07K 7/04; A23J 1/14; A61K 35/78; C12N 15/00, 9/12
C07H 15/12

II. FIELDS SEARCHED

Minimum Documentation Searched 4

Classification System	Classification Symbols
U.S.	530/350, 377, 825; 435/172.3, 194; 536/27

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched 6

COMPUTER SEARCH: CAS, BIOS
DNA OR TOXIN WITH LIGAND OR BINDING PROTEIN AND CYSTEINE

III. DOCUMENTS CONSIDERED TO BE RELEVANT 14

Category *	Citation of Document, 16 with indication, where appropriate, of the relevant passages 17	Relevant to Claim No. 18
X	Eur. J. Biochem. Vol. 134 issued August 1983, (Berlin Germany), (MAASSEN ET AL), "Synthesis and Application of Two Reagents for the Introduction of Sulfhydryl Groups into Proteins." pages 327-330, especially page 327.	1-14
X	WO,A, 83/03971 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 24 November 1983 (24.11.83), see pages 2 and 7.	1-14
X	The EMBO Journal, Vol. 4 issued 1985 (Oxford England) (JACOB ET AL), "Priming immunization against cholera toxin and E.coli heat-labile toxin by a cholera toxin short peptide-β galactosidase hybrid synthesized in E. coli," pages 3339-3343.	1-14

* Special categories of cited documents: 15

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search 2	Date of Mailing of this International Search Report 2
12 January 1987	09 FEB 1987
International Searching Authority 1	**Signature of Authorized Officer 20**
ISA/US	ALVIN E. TANENHOLTZ

Form PCT/ISA/210 (second sheet) (October 1981)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No [18]
X	Journal of Cellular Biochemistry Vol. 20, issued 20 August 1982 (New York, N.Y. USA) (HERSCHMANN ET AL), "Toxic Ligand Conjugates as Tools in the Study of Receptor-Ligand Interactions", pages 163-176.	1-14
X	The Journal of Biological Chemistry, Vol. 258 issued 10 February 1983 (Baltimore Maryland USA) (BACHA ET AL), "Thyrotropin-releasing Hormone-Diptheria Toxin-releated Polypeptide Conjugates", pages 1565-1570.	1-14
X	WO,A, 84/00299 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 2 February 1984 (02.02.84), see pages 1-4	1-14
X	US,A, 4,468,382 (BACHA ET AL) Published 28 August 1984	1-14
X	VICTOR J. HARUBY ET AL, "Peptides, Structures and Function: Proc. of the Eighth American Peptide Symp." Published 1983 by Pierce Chemical Company (Rockford Illinois USA), see pages 837-852	1-14

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

X	Proc. Natl. Acad. Sci. USA Vol. 80 issued November 1983 (Washington D.C.), (GREENFIELD ET AL), "Nucleotide sequence of the structural gene for diptheria toxin carried by corynebacteriophage β," pages 6853-6857.	1-14
X	The Journal of Biological Chemistry, Vol. 256 issued 10 June 1981 (Baltimore Maryland USA), (ROTH ET AL), "Insulin-Ricin B Chain Conjugate", pages 5350-5354	1-14

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [10]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers ..... because they relate to subject matter [12] not required to be searched by this Authority, namely:

2. ☐ Claim numbers ....., because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [13], specifically:

VI. ☐ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [11]

This International Searching Authority found multiple inventions in this international application as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.